

## EL DISCURSO

**Barbara J. Grosz, Martha E. Pollack y Candace L. Sidner**

La investigación en análisis del discurso, estructura del discurso y procesamiento del discurso considera a los discursos, no a las oraciones aisladas, como las unidades básicas del lenguaje. En todos los empeños es central el deseo de descubrir los mecanismos que hacen que ciertas secuencias de preferencias sean coherentes y otras no y de comprender cómo afecta el contexto al uso y a la comprensión de varias clases de expresiones lingüísticas. En este capítulo nos centramos en tres grupos de problemas del discurso: aquéllos que surgen al tratar de determinar la estructura del discurso, aquéllos relacionados con fenómenos de nivel sintagmático (por ejemplo, pronombres, descripciones definidas, elipsis), y aquéllos que conciernen a cómo se deriva el significado o efecto de un discurso o segmento discursivo. Dado que en las ciencias cognitivas se ha hecho más investigación de la que puede recogerse en este capítulo, nos ocupamos aquí del trabajo que ha tomado como centrales las cuestiones computacionales o de procesamiento.

Comenzamos con un repaso de las primeras aproximaciones al discurso en inteligencia artificial. Estas primeras aproximaciones hicieron evidentes los problemas de las teorías existentes, la necesidad de determinar el papel de la información no lingüística (incluyendo conocimiento acerca del dominio, diferencias entre las creencias de hablante y oyente, lector y escritor, y estado cognitivo) en el procesamiento del discurso, y la necesidad de teorías adecuadas, computacionalmente tratables, para sustentar el procesamiento del discurso. A continuación describimos las aproximaciones a la estructura del discurso, discutiendo varias ideas sobre las clases de segmentos discursivos que existen y las clases de relaciones que los unen. En la línea de trabajos anteriores (Grosz y Sidner 1986), distinguimos la estructura lingüística del discurso de su estructura intencional y de su estado atencional cambiante.

En la sección 11.3 examinamos diversas clases de información sobre la estructura del discurso que se aporta explícitamente en la propia señal lingüística. En la sección 11.4 describimos la investigación acerca de los efectos de la estructura y contexto discursivos sobre la interpretación de varios fenómenos de nivel sintagmático, incluidas las descripciones definidas, los pronombres, los sintagmas nominales

cuantificacionales (por ejemplo, *todas las chicas del bloque*), la referencia de eventos, y la elipsis. La sección 11.5 describe varias formulaciones de planes y técnicas que se han desarrollado para el reconocimiento de planes. Una vez que tomamos seriamente en cuenta la idea de que el lenguaje es usado por los agentes para influir en su mundo, se torna crucial comprender los efectos que una preferencia individual puede tener (éste era el alcance inicial de la teoría de los actos de habla), y los modos en los que las preferencias pueden combinarse para formar acciones mayores (investigación de reconocimiento de planes en el discurso). Así las acciones lingüísticas pueden verse como acciones que influyen en la estructura y el significado discursivos.

### **11.1 Primeras Aproximaciones al Procesamiento del Discurso**

Los primeros intentos de tratar directamente los problemas computacionales del discurso se realizaron en los primeros años 70 en el contexto de la construcción de sistemas completos de procesamiento del lenguaje natural basados en el computador. Estos sistemas no sostenían ningún modelo de la estructura discursiva tal y como se entiende ahora. Sin embargo, los creadores de sistemas experimentaron con técnicas de comprensión de pronombres y una forma limitada de reconocimiento de intenciones. El sistema LUNAR (Woods et al. 1972, Woods 1978) proporcionaba mecanismos de resolución de pronombres, pero dado que cada par de pregunta-y-respuesta era un intercambio independiente, el modelo del discurso era simple: sólo se mantenía una lista de entidades mencionadas previamente para usarla en la comprensión de pronombres. Charniak (1977) intentó identificar los referentes de las descripciones definidas y los pronombres en el proceso de "comprensión" de relatos infantiles codificando la información sobre el dominio en reglas de inferencia y programas desencadenantes asociados. El sistema de diálogo más extenso, el sistema SHRDLU de Winograd (Winograd 1971, 1972) participaba con un usuario en un diálogo sobre construcciones de bloques de juguetes. SHRDLU interpretaba las intenciones del usuario como programas de realizar acciones y conservaba una historia de las acciones que realizaba. Podía interpretar algunos pronombres personales, algunas descripciones definidas y expresiones anafóricas de tipo "el/la que" usando "especialistas de palabras", es decir, programas que determinaban la referencia basados en heurísticos de plausibilidad para los objetos mencionados más recientemente en la oración previa.

La siguiente generación de sistemas de lenguaje natural, de finales de los 70, se concentraba en el uso de conocimiento acerca del dominio en el discurso o bien en métodos de incorporación de otros fenómenos discursivos aparte de la comprensión de pronombres. La mayoría de los sistemas limitaron notablemente el dominio del discurso a una pequeña clase de acciones o eventos y se centraron en interacciones que se ceñían a una estrecha gama de comportamiento discursivo (por ejemplo, SAM (Cullingford 1977, 1981) y GUS (Bobrow et al. 1977)). Sin embargo, Lehnert (1977) experimentó con un sistema para captar el significado no-literal de preguntas planteadas a un programa de comprensión de relatos. Grosz y sus colaboradores (Grosz 1977, Walker 1978) desarrollaron el Sistema de Comprensión de Diálogos Tarea, en el que se distinguió entre conocimiento del dominio (representado en un modelo de la tarea), información del discurso (representado en un foco global y algoritmos asociados para la interpretación de referencias), y reconocimiento de intenciones (tratamiento de preguntas sobre, y adaptaciones del, modelo de la tarea). Este sistema fue el primero en considerar la interacción entre la estructura (lingüística) de un discurso y la interpretación de expresiones en el discurso. No distinguió con suficiente claridad, sin embargo, los diferentes papeles de un modelo de tarea en el procesamiento del discurso.

Las primeras aproximaciones al procesamiento del discurso enseñaron a los investigadores que el procesamiento del discurso no podía verse como un aglomerado de procedimientos para la comprensión de referencias, reconocimiento de intenciones y manipulación de conocimiento del dominio. Se requiere un medio para distinguir entre estos procesos y proporcionar mecanismos para cada uno así como para su interacción.

## **11. La Estructura del Discurso**

Las preferencias que conforman un discurso no son secuencias arbitrarias, sino que tienen estructura de modo similar al que las palabras en una oración individual tienen estructura sintáctica. El análisis de una variedad de tipos de discurso ha establecido que los discursos se dividen en segmentos discursivos y que éstos pueden tener diferentes relaciones entre ellos. Entre los tipos de discurso que han sido analizados se encuentran diálogos orientados a tarea (ver, por ejemplo, Grosz 1978a, Mann et al. 1975, Sidner 1982), descripciones de objetos complejos (Linde 1979), narraciones (Polanyi 1985, Schiffrin

1982), argumentos (tanto informales (Reichman-Adar 1984) como formales (Cohen 1984)), negociaciones (Linde y Goguen 1978), y explicaciones (Reichman-Adar 1984).

La comprensión de la estructura del discurso es importante tanto para las teorías del significado del discurso como para el procesamiento del lenguaje. Las teorías del significado del discurso dependen en parte de la especificación de las unidades básicas del discurso y de las relaciones que pueden darse entre ellas. El procesamiento del discurso requiere habilidad para determinar con qué partes del discurso se relaciona una preferencia individual. Así el papel de la estructura en el procesamiento del discurso se deriva tanto de su papel en la delimitación de las unidades de significado discursivo como de la restricción de las unidades de discurso que se consideren pertinentes para la interpretación de cualquier preferencia individual.

Además, la explicación de la interpretación de una variedad de expresiones lingüísticas depende de la explicación del papel de algunas de estas expresiones al determinar la estructura del discurso así como de la comprensión del efecto de la estructura del discurso al restringir la interpretación de otros. La sección 11.3 discute el uso de sintagmas de entrada [*cue phrases*] y de la entonación para marcar la estructura del discurso. La sección 11.4 describe la investigación sobre la interacción entre la estructura del discurso y la interpretación de expresiones anafóricas.

Aunque hay un acuerdo general sobre el hecho de que los discursos están compuestos por segmentos y sobre la interacción entre la estructura del discurso y la interpretación de varias clases de expresiones lingüísticas, las teorías de la estructura del discurso postulan diferentes tipos de información como centrales para la computación de la estructura del discurso. Además las teorías difieren en las clases de relaciones entre segmentos consideradas importantes. Los primeros trabajos en comprensión de texto (vanDijk 1972, Rumelhart 1975) proponían gramáticas de texto (o relato) análogas a las gramáticas oracionales<sup>1</sup>; el primer trabajo sobre diálogo (Grosz 1974) sostenía que los diálogos orientados a tarea tenían una estructura que dependía de la estructura de la tarea que se estaba realizando. El trabajo posterior ha tomado una de las siguientes aproximaciones: (1)

---

<sup>1</sup>Las gramáticas de relatos como base para el procesamiento han sido extensamente criticadas; el artículo de Levy (1979) sobre la estructura del discurso contiene un buen sumario.

adaptaciones de la noción de gramática, (2) especificación de un pequeño conjunto de relaciones retóricas o textuales como la base de la estructura del discurso, (3) investigación del conocimiento específico del dominio o del sentido común como fuente de la estructura del discurso, o (4) examen de las intenciones interpretadas ampliamente (una generalización de la estructura de la tarea) y las relaciones entre ellas como base de la estructura del discurso.

### **Segmentación y Estructura de las Preferencias**

La investigación sobre las relaciones posibles entre las preferencias que comprende un discurso ha considerado tales cuestiones como qué preferencias se agrupan en un segmento y qué relaciones pueden darse entre estos segmentos (Polanyi 1985, Hobbs 1979, Linde 1979, Mann y Thompson 1986). En un trabajo previo (Grosz y Sidner 1986) nos hemos referidos a esta estructuración del discurso en grupos como la estructura lingüística del discurso. Nuestra terminología refleja el acento puesto a la estructura del propio comportamiento lingüístico. En las teorías que consideran primariamente la estructura lingüística, la estructura discursiva se representa típicamente como un árbol cuyos nodos son preferencias de nivel de cláusulas o grupos de ellas<sup>2</sup>.

Las teorías difieren con respecto a la información sobre la que basan sus decisiones acerca de las relaciones de incrustación codificadas en el árbol y con respecto a las clases específicas de relaciones de incrustación que consideran posibles. Aunque algunas diferencias son resultado de los tipos de discursos tratados (por ejemplo, narraciones frente a argumentos), otras reflejan la disparidad de tesis sobre los objetivos del procesamiento del discurso.

Linde (1979) y Polanyi (1986) proponen modelos que en parte derivan de la tradición sociolingüística, en la que las explicaciones se expresan sólo en términos de comportamiento superficial. Así se ocupan de proporcionar una explicación de la estructura del discurso en términos de las conexiones superficiales que se obtienen entre las preferencias. Polanyi sostiene que la estructura jerárquica del discurso "emerge de las

---

<sup>2</sup>Se considera al árbol como análogo a los árboles analíticos para las oraciones individuales. En un artículo anterior (Grosz y Sidner 1986) sostenemos que las relaciones de incrustación (o constituencia) representadas en el árbol se derivan del comportamiento del apilamiento del estado atencional.

relaciones estructurales y semánticas que se obtienen entre las unidades lingüísticas que los hablantes usan para construir sus discursos" (Polanyi 1986, p. 4). En su modelo un árbol discursivo se construye usando un conjunto de gramáticas discursivas. Los nodos del árbol son cláusulas (en algunos casos con su semántica asociada). Es decir, la estructura del discurso se toma como una estructura arbórea con elementos lingüísticos como nodos.

La investigación en este paradigma trata normalmente a las interacciones completas (por ejemplo, conversaciones personales extensas, interacciones doctor-paciente) como las mayores unidades de discurso, y a las cláusulas, quizá agrupadas por relaciones de coordinación o subordinación (por ejemplo, listas, expansiones) como las menores unidades. Las unidades intermedias de explicación pueden ser motivadas socialmente (por ejemplo, encuentros de servicio), exámenes médicos) o motivadas lingüísticamente (por ejemplo, relatos). Frecuentemente se adjudican etiquetas a estos niveles (por ejemplo, pregunta/respuesta) que indican las intenciones de los participantes en el discurso, pero las teorías evitan cualquier referencia a la intención en sus explicaciones.

### **Relaciones Retóricas**

Reichman-Adar (1984) estudió argumentos informales; la estructura que propuso se basa en una gramática para argumentos análoga a las gramáticas de relato o de texto (van Dijk 1972, Rumelhart 1975). Por ejemplo, sostiene que un argumento consiste en una afirmación seguida de un respaldo, y que el respaldo se descompone en una cita de autoridad o una narración o una proposición. Aunque las relaciones que postula entre las preferencias y los segmentos discursivos combinan propiedades semánticas, pragmáticas e intencionales, el análisis del procesamiento de Reichman se centra en el nivel lingüístico.

Otras investigaciones que construyen una estructura de árbol con nodos lingüísticos se ha centrado en las relaciones que pueden darse entre los significados de las preferencias (es decir, en relaciones semánticas). Cohen (1979) ha analizado una variedad de argumentos e investigado el problema de inferir relaciones de evidencia entre las proposiciones expresadas en ellos. Presenta un modelo de procesamiento de los argumentos que utiliza una combinación de información de sintagmas de entrada y (un conjunto inicial de) métodos para inferir relaciones de evidencia para determinar la estructura del argumento.

Buena parte de la investigación descrita hasta ahora, así como los trabajos de Hobbs (1979), Lehnert (1981), Mann y Thompson (1986) y McKeown (1985) ofrecen un estrato subyacente de relaciones retóricas como base para derivar relaciones entre las preferencias y los segmentos. Cada una de estas relaciones requiere el procesamiento de información del dominio para determinar cómo se puede reconocer o producir una relación en el procesamiento del discurso. Por ejemplo, Hobbs (1979) define un conjunto de relaciones de coherencia (por ejemplo, paralelo, capacitación, contraste; se puede encontrar una lista completa en Hobbs (1983)) que se dan entre los segmentos discursivos. La determinación de la relación que se da entre dos preferencias sucesivas depende de las inferencias realizadas sobre la base de hechos del dominio. Las relaciones de Lehnert se derivan de un modelo subyacente de estados mentales y eventos. Un conjunto primitivo de configuraciones, que denomina *unidades de trama*, describe las transiciones permisibles entre estados o eventos y por ello codifica directamente la información del dominio. Estas unidades de trama funcionan entonces como las relaciones entre las preferencias de los segmentos. En el trabajo sobre generación de discursos de McKeown, se organiza una colección de relaciones retóricas (que incluye, por ejemplo, atribución, identificación y comparación) en esquemas que definen abstractamente el conjunto de tipos de discurso aceptables; se produce al menos una preferencia (o más) como ejemplar(es) de los predicados en un esquema.

### **Intenciones y Atención en la Estructura del Discurso**

Levy (1979) y Grosz y Sidner (1986) han sostenido que las relaciones basadas en planes o relaciones intencionales son la raíz de la estructura discursiva y postulan sólo una relación de incrustación en la estructura lingüística. Las relaciones de incrustación entre segmentos dependen en parte de ciertas características lingüísticas del segmento (como la prosodia o los sintagmas de entrada), y en parte de las intenciones transmitidas por las preferencias del segmento. La intención de nivel segmento, sin embargo, no es una mera función de las intenciones de nivel de preferencia, sino una función compleja de preferencias, hechos del dominio, intenciones de nivel de preferencia, e inferencias sobre éstas.

Nosotras sostenemos también (Grosz y Sidner 1986) que la estructura del discurso es un compuesto de tres estructuras interrelacionadas; además de la estructura lingüística hay una

estructura de intenciones y un estado atencional. La estructura intencional comprende propósitos de segmento discursivo y relaciones entre ellos. Los propósitos de segmento discursivo son intenciones de los participantes en el discurso que conducen en parte al discurso y se pretende, como las intenciones de nivel de preferencia de Grice (1957, 1968), que sean reconocidas. Al contrario que las relaciones retóricas discutidas previamente, los propósitos de segmento discursivo no se derivan de un conjunto particular de intenciones; casi todo lo que pueda ser objeto de una intención puede ser un propósito de segmento discursivo. Sin embargo, se definen dos relaciones entre intenciones como comunes a muchos discursos: *dominio* y *precedencia de satisfacción*. Estas relaciones representan, respectivamente, el hecho de que la satisfacción de una intención contribuya a la satisfacción de otra, y el hecho de que una intención deba satisfacerse antes que otra. El reconocimiento de estas dos relaciones desempeña en nuestra teoría el mismo papel que el reconocimiento de las relaciones retóricas en las teorías descritas previamente. La determinación de los propósitos de segmento discursivo depende en parte del reconocimiento de estas relaciones, en parte del conocimiento del dominio, y en parte de otras características del contexto discursivo.

El estado atencional refleja el foco de atención de los participantes del discurso según avanza el mismo. Se modela como una pila (pushdown) de espacios focales, uno por cada segmento del discurso. Los espacios focales contienen representaciones del propósito del segmento discursivo y de las entidades referidas en el segmento. La pila crece cuando se introducen los segmentos en un discurso y decrece a medida que se satisfacen las intenciones de los segmentos. De acuerdo con la terminología estándar, decimos que un nuevo espacio se *empuja* a la pila cuando se introduce un nuevo segmento, y que un espacio *salta* cuando se satisface su propósito. El modelo de espacios focales del estado atencional constriñe el procesamiento según avanza el discurso. Esas entidades y propósitos representados en la pila son los más prominentes del discurso. Las restricciones pueden utilizar este hecho para estipular cuándo pueden usarse varias expresiones lingüísticas y para ayudar a determinar cuándo un propósito dado de segmento discursivo puede dominar o preceder en la satisfacción a otro.

### 11.3 Indicadores Lingüísticos de la Estructura Discursiva

La moraleja de este capítulo debería estar ahora clara: los discursos exhiben estructura. Todo discurso puede ser dividido intuitivamente en segmentos, y los segmentos pueden tener incrustados a, o incrustarse en, otros, resultando en una estructura jerárquica. Reconocer la estructura de un discurso es parte esencial de su comprensión.

¿Pero cómo se reconoce la estructura de un discurso? Como hemos señalado en la sección precedente, la respuesta a esta pregunta es todavía ampliamente debatida—lo que no es sorprendente, puesto que es la cuestión motivadora central en la investigación sobre el discurso. Virtualmente todos los investigadores están de acuerdo, no obstante, en que los hablantes disponen de ciertos mecanismos lingüísticos poderosos para ayudar a los oyentes a reconocer la estructura del discurso en curso. Estos mecanismos, tales como los sintagmas de entrada, los patrones prosódicos y los gestos, pueden indicar dónde comienza o termina un segmento discursivo y cómo está relacionado con otros segmentos.

### **Sintagmas de entrada**

La clase de indicadores lingüísticos de estructura discursiva más estudiada es la de los *sintagmas de entrada* (a veces denominados *palabras indicio*, *marcadores discursivos* o *partículas discursivas*). Los sintagmas de entrada son expresiones como *ahora bien*, *en primer lugar* y *dicho sea de paso*, que no realizan una contribución semántica directa a una preferencia, sino que aportan información sobre la estructura del discurso que contiene a la preferencia. Por supuesto que dado que las diversas teorías sostienen tesis diferentes acerca de lo que supone la estructura de un discurso, existen diferencias entre las teorías sobre exactamente qué información señala un sintagma de entrada.

Los investigadores que ven la estructura de un discurso como compuesta solamente por una estructura lingüística sostienen que los sintagmas de entrada proporcionan una indicación directa de cómo se relacionan los diversos segmentos en la jerarquía (Cohen 1984). Aquéllos que argumentan que la estructura de un discurso consiste en un estrato subyacente de relaciones retóricas conectando los segmentos discursivos ven los sintagmas de entrada sugiriendo la relación particular existente entre dos segmentos discursivos. Por ejemplo, Hobbs (1985, p. 31) afirma que "'Estos es' o 'i.e.' sugiere *elaboración*, 'de modo similar' sugiere *paralelo*, 'por ejemplo' sugiere *ejemplificación* y

'pero' sugiere *contraste* o *expectativa violada*"<sup>3</sup>. Afirmaciones semejantes se han hecho en la Teoría de la Estructura Retórica (Mann y Thompson 1986) y en explicaciones de análisis sintáctico del discurso (Polanyi y Scha 1984, Reichman-Adar 1984).

Grosz y Sidner (1986) sostienen una tesis sobre el papel de los sintagmas de entrada que contrasta con las anteriores. Como se ha indicado antes, hemos argumentado que la estructura de un discurso en realidad comprende tres estructuras interrelacionadas. En esta perspectiva los sintagmas de entrada pueden proporcionar información sobre uno o más componentes de la estructura (tripartita) global (Grosz y Sidner 1986, pp. 196-199). Por ejemplo, ciertos sintagmas de entrada, tales como *eso me recuerda* o *de todos modos*, indican cambios en el estado atencional—el primero indica un empujón a un nuevo espacio focal y el segundo un salto a un espacio establecido previamente. Nótese que aunque al cambio de estado atencional señalado por *eso me recuerda* le acompaña la adición de nuevos componentes a la estructura intencional, el sintagma de entrada por sí solo no especifica de qué cambio se trata. Con *de todos modos* no hay ningún cambio de estructura intencional. La expresión *a propósito* proporciona la información de que el hablante va a emprender una digresión y que, en consecuencia, se va a extender la estructura intencional con una nueva jerarquía intencional distinta a la existente. La expresión *por ejemplo* indica que la intención subyacente al segmento del discurso siguiente—aquél cuyo comienzo se marca mediante el segmento de entrada—está dominada en la estructura intencional por la intención subyacente al segmento del discurso anterior.

La mayoría de los investigadores, independientemente de su postura sobre lo que constituye realmente el discurso, creen que los sintagmas de entrada no son en general ni necesarios ni suficientes para determinar la estructura del discurso. Hay muchos casos en los que es bien posible determinar la estructura, de un discurso, o una parte, que carece de sintagmas de entrada. Igualmente, hay muchos discursos, o partes de éstos, que contienen sintagmas de entrada que sólo sugieren la estructura subyacente o que, dicho de otro modo, proporcionan restricciones sobre la serie de estructuras posibles. Por último, la estructura de un discurso depende de la información

---

<sup>3</sup>Hobbs (1985) sostiene que los hablantes no sólo usan los sintagmas de entrada para señalar la estructura discursiva, sino que tratar de insertarlos en un discurso completo puede ser útil para el analista que trata de discernir su estructura.

proporcionada por las preferencias que contiene y el modo en el que esa información está interconectada. Los sintagmas de entrada simplifican la labor de determinar estas interconexiones. Se ha mostrado que el proceso de determinación de las relaciones semánticas interoracionales (Cohen 1984) y de reconocimiento de planes (Litman and Allen 1988) puede ser restringido teniendo en cuenta los sintagmas de entrada.

### **Prosodia y Gesticulación.**

La prosodia es otro instrumento eficaz para señalar la estructura del discurso. Los estudios de discursos producidos espontáneamente han mostrado que los cambios en la longitud de pausa (Chafe 1979, 1980) y en la velocidad del habla (Butterworth 1975) están correlacionados con los límites de los segmentos del discurso. Hirschberg, Pierrehumbert, Litman y Ward han proporcionado datos de una estrecha correlación entre ciertas características prosódicas y componentes específicos de la estructura discursiva tripartita propuesta por Grosz y Sidner (Hirschberg et al. 1987, Hirschberg y Pierrehumbert 1986). El *registro tonal* puede señalar los límites de los segmentos del discurso (Silverman (1987) llegó a la misma conclusión); la *acentuación* puede proporcionar información sobre el estado atencional, y el *tono* puede marcar la estructura intencional. Además, el estilo y la acentuación pueden ayudar a distinguir el uso de una expresión como *ahora* como sintagma de entrada del uso en que la expresión contribuye directamente al significado oracional (Hirschberg y Litman 1987).

La gesticulación es otro instrumento útil, al menos en las conversaciones cara a cara, para señalar la estructura discursiva. Hay estudios que muestran que la gesticulación tiende a coincidir con los límites de los segmentos discursivos (Kendon 1972, Marslen-Wilson et al. 1982) y puede proporcionar información sobre el foco de atención y la estructura intencional (McNeill 1979, McNeill and Levy 1982).

### **11.4 Fenómenos de nivel sintagmático.**

El contexto discursivo tiene efectos sobre la interpretación de sintagmas individuales dentro de una sola preferencia. El significado de pronombres y descripciones definidas se ve bastante obviamente influenciado por el contexto en el que son usados; los problemas relativos a la influencia del contexto sobre los procesos de generación e interpretación de tales sintagmas se

encuentran entre los más estudiados en procesamiento del discurso. También ha sido abordado el problema, relacionado con los anteriores, de representar apropiadamente los sintagmas nominales cuantificacionales. En la subsección siguiente discutiremos una serie de aproximaciones a estos problemas de referencia de sintagmas nominales. La referencia puede ser a objetos así como a eventos, lo que da lugar a problemas a discutir en la subsección segunda. La adjunción de modificadores o la elipsis discursiva son otros problemas de nivel sintagmático que han sido tratados desde la perspectiva del procesamiento del discurso.

De los tres componentes de la estructura del discurso, el estado atencional es el que tiene mayor efecto sobre los problemas de interpretación en el nivel sintagmático. La investigación sobre expresiones referenciales es la que más directamente ha usado modelos atencionales, mientras que las aproximaciones a la adjunción de modificadores y a la elipsis discursiva no han hecho más que empezar a usarlos. Aunque todavía no hay una descripción completa del conjunto de las restricciones discursivas (ni de las perceptuales) relevantes para ninguno de estos problemas de nivel sintagmático, los investigadores han identificado y experimentado con una serie de representaciones y algoritmos centrales para éstas.

### **Pronombres y Descripciones Definidas.**

La preocupación principal de las teorías y modelos computacionales del uso de expresiones referenciales ha sido determinar la entidad (o entidades) a la que refiere un pronombre o una descripción definida<sup>4</sup>. Se ha investigado tanto el problema de especificar la serie de posibles referentes que una primera descripción hace disponible como el de elegir entre las distintas posibilidades.

Para el problema de identificar el referente de un sintagma anafórico<sup>5</sup> se han adoptado dos aproximaciones. En una aproximación (Hobbs 1979) la identificación del referente se

---

<sup>4</sup>Este énfasis difiere del de gran parte de la lingüística americana, en la que han sido centrales las restricciones que impiden la co-especificación de pronombres (véase, por ejemplo, la investigación sobre anáfora pronominal de Reinhart (1983) y Lasnik (1976)).

<sup>5</sup>Usamos aquí *anáfora* y *anafórico* para referirnos a aquellos sintagmas que refieren a entidades que han sido mencionadas previamente en el discurso y que co-especifican (compárese con Sidner 1983) junto con algún sintagma que aparece previamente en el discurso.

subsume a procesos más generales de inferencia. En la otra aproximación se consideran primarias las cuestiones de cómo interactúan las expresiones referenciales con el estado atencional (Grosz 1977, Sidner 1981, Reichman-Adar 1984, Grosz and Sidner 1986). La investigación en la primera aproximación sólo considera problemas de interpretación de las expresiones referenciales, mientras que la investigación dentro de la segunda también se preocupa de especificar restricciones sobre la generación de expresiones referenciales adecuadas.

Para la segunda aproximación los conceptos de foco y proceso de focalización (Grosz 1977, Grosz 1978b, Grosz 1981, Grosz and Sidner 1986) han desempeñado papeles centrales en el tratamiento de las descripciones definidas, usadas para primera referencia o como sintagmas nominales anafóricos. En este trabajo la focalización se define como el movimiento del foco de atención de los participantes en el discurso según éste avanza. Se han identificado dos niveles de focalización, global y local (Grosz 1977, Grosz and Sidner 1986). La focalización global se modela mediante una pila de espacios focales; afecta al uso e interpretación de descripciones definidas. La focalización a nivel local se modela con centros y centralización. El centro de un segmento dado es un elemento del estado atencional y al comienzo de cada nuevo segmento se introduce un nuevo centro. La centralización afecta al uso y la interpretación de pronombres.

Cada espacio individual de la pila focal global contiene representaciones de las entidades enfocadas por los participantes durante cierto segmento del discurso así como el propósito del segmento discursivo. Las entidades actualmente en foco (esto es, en algún espacio de la pila focal) son las primeras candidatas para ser referentes de las descripciones definidas; son también el origen de las entidades enfocadas implícitamente (esto es, un sintagma puede referir a un ítem relacionado con algo de un espacio focal actual; por ejemplo, *la portada* puede usarse para referirse a la portada de un libro cuando ese libro está en el foco). El conjunto de entidades en el foco global proporciona también restricciones sobre el contenido de las descripciones definidas subsiguientes; por ejemplo, un hablante debe incluir suficientes descriptores para distinguir la entidad a la que se quiere referir de otras entidades en el foco.

Grosz (1977) ideó un conjunto de mecanismos focalizadores para la interpretación de sintagmas nominales definidos en un sistema que participaba en un diálogo sobre una

tarea.<sup>6</sup> Los mecanismos traían entidades al foco a medida que el discurso se desplazaba a una subtarea de la tarea general y las retiraban del foco global cuando se completaba la subtarea. La posición en el foco global era dependiente de la subtarea más inmediata; las otras subtareas incompletas y sus entidades asociadas estaban en el foco global, pero eran menos prominentes que aquéllas asociadas a la subtarea más inmediata. Los procedimientos para la interpretación de sintagmas nominales definidos y anafóricos y de las primeras referencias enfocadas implícitamente eligen los ítems entre aquéllos que están en el foco global. Las técnicas de focalización podían predecir el referente anafórico de un sintagma nominal definido como *el tornillo* cuando se ha traído al foco el tornillo de la polea. Limitaban también la búsqueda necesaria para hallar la representación de una entidad mencionada previamente—aquellas entidades que no estaban en el foco no estaban disponibles, y aquéllas en el foco y asociadas a la subtarea más inmediata se buscaban antes que las menos prominentes. Por ejemplo, *el tornillo* sería entendido como un referente del tornillo de la polea sólo cuando la polea fuera más prominente que algún otro tornillo (por ejemplo, el tornillo del taladro) en el foco global; en caso contrario se elegiría como referente al tornillo del taladro.

Appelt y Kronfeld (1987) han utilizado los mecanismos de focalización para la generación de expresiones referenciales de primer uso. Además Kronfeld (1986) ha proporcionado los conceptos de descripciones funcionalmente pertinentes (esto es, aquéllas cuyo contenido se requiere para distinguir el referente) y descripciones conversacionalmente pertinentes (esto es, aquéllas cuyo contenido proporciona información sobre la relación del referente con otros aspectos de la conversación). Estos conceptos redefinen la distinción clásica de Donnellan (1966) entre descripciones referenciales y atributivas en términos del papel que un sintagma desempeña como instrumento de referencia y como contribuyente al contenido del discurso.

El foco inmediato (Sidner 1979) guía la interpretación y la generación de los pronombres de tercera persona así como los usos anafóricos de *esto* (*éste*, *ésta*) y *eso* (*ése*, *ésa*). La focalización inmediata opera dentro de los segmentos discursivos individuales; localiza la entidad más relevante en cualquier

---

<sup>6</sup> Para este sistema los propósitos de segmento discursivos eran tratados como tareas y subtareas.

preferencia dentro del segmento, basándose en características del contexto precedente, particularmente en propiedades de la preferencia precedente. Sidner (1979, 1981, 1983) desarrolló un conjunto de algoritmos para predecir la elección de focos locales en el discurso y su desplazamiento a medida que el discurso avanzaba. Sus algoritmos hacían uso de los conceptos de un foco discursivo y un foco actor, es decir, entidades mencionadas en la preferencia que estaban localmente en foco debido a la estructura sintáctica y a las relaciones temáticas de la preferencia. Un conjunto de reglas, que utilizaban los focos inmediatos y un conjunto de nuevos focos potenciales, predecía la interpretación intendida de pronombres en preferencias subsiguientes. McKeown (1985) y McDonald (1983) adaptaron los algoritmos de focalización para generar pronombres en el texto.

La teoría de centrado de Grosz, Joshi y Weinstein (Joshi y Weinstein 1981, Grosz et al. 1983) reemplazó la noción de focos inmediatos por la de centros: un centro de mirada-hacia-atrás (Cb) que corresponde aproximadamente al foco del discurso y varios centros de mirada-hacia-adelante (Cf) que corresponden a los nuevos focos potenciales. Como en la teoría de Sidner los centros cambian de acuerdo con el comportamiento de las expresiones referenciales en cada preferencia del discurso. La teoría difiere en que no tiene correlato alguno del foco actor y en que da cuenta de los usos múltiples de un mismo pronombre permitiendo que se pronominalicen otras entidades distintas de Cb siempre que ésta también lo esté. La teoría de centrado simplifica la explicación de Sidner de la focalización local eliminando la necesidad del foco del discurso así como del foco actor y proporcionando una explicación de ciertas interpretaciones de pronombres que resultan problemáticas en la teoría de los focos inmediatos.

Aunque el foco de atención es crucial para comprender la anáfora, la forma y el contenido de los sintagmas y preferencias que primero evocan las entidades a las que refiere la anáfora también contribuye vitalmente a su comprensión. Webber (1980, 1983) introdujo la noción de sintagmas “que evocan entidades discursivas”. Con esto quiere decir que los sintagmas traían al discurso, o *evocaban naturalmente*, una colección bien estructurada de representaciones de las entidades (reales) referidas. Para la interpretación de la anáfora definida y la elipsis de sintagma verbal estaban disponibles las entidades discursivas. Definió el concepto de *descripción invocadora de entidad discursiva* (ID) para formalizar su noción y postuló un conjunto

de representaciones y reglas específicas para crear estas IDs por la aplicación sucesiva de las reglas.

Las reglas de Webber operaban sobre representaciones semánticas complejas, que incluían cuantificadores incrustados y cuantificación sobre conjuntos de individuos. Las reglas producían IDs para cada posible interpretación de los sintagmas nominales y sintagmas verbales de una preferencia, pero Webber no exploró cómo determinar computacionalmente qué interpretación era la intendida sobre la base de descripciones anafóricas subsiguientes. Sidner (1983), sin embargo, esbozó una explicación de cómo la focalización podría proporcionar la información necesaria. En trabajos posteriores Kamp (1981) y Heim (1982) han propuesto formalismos alternativos semejantes; Guenthner y sus colaboradores (1986) describen un sistema que utiliza el formalismo de Kamp. Dos intentos de construcción de sistemas han investigado métodos de para computar las representaciones de sintagmas nominales complejos como parte del proceso de determinar (una representación de) el significado de una oración (Dahl 1987, Pollack y Pereira 1988).

Entre las expresiones referenciales menos investigadas en la investigación computacional del discurso se encuentra el uso de la deixis<sup>7</sup>. Fillmore (1975) subcategorizó cuatro tipos de deixis: espacial, temporal, social y discursiva. La estructura del discurso puede restringir los dos últimos tipos como se ejemplifica por sintagmas como *los primeros*, usado para referirse a los dos primeros elementos de una lista (una instancia de deixis discursiva), y *nosotros* (una instancia de discurso social). Sidner (1979) exploró algunos casos de deixis discursiva, pero queda por hacer un tratamiento general.

## La referencia a eventos

La investigación del discurso se ha ocupado de dos problemas causados por la referencia a eventos: identificar el evento al que se refiere una referencia anafórica y determinar el relativo orden temporal de los eventos descritos en el discurso.

Las referencias anafóricas usan típicamente formas pronominales tanto puras (*lo, esto, eso*) como en el ejemplo 1, o

---

<sup>7</sup> *Deixis*, del griego para “apuntar”, se usa para referirse a sintagmas demostrativos (por ejemplo, *este libro* y *eso*) y a otros constructos lingüísticos cuya interpretación de modo semejante requieren apuntar a alguna circunstancia de la preferencia (por ejemplo, *aquí* y *ahora*).

con el verbo hacer (*hace así, lo hace, hace eso*), como en el ejemplo 2. La forma con *hacer* y el pronombre puro parecen comportarse de manera diferente. Aunque hay algunas investigaciones sobre las formas con *hacer* (Sidner 1979, Sidner 1981, Robinson 1981), queda por hacer una explicación detallada de todas las referencias directamente anafóricas a eventos.

- (1) Juan corre todos los días de la semana.  
*Ese* es su principal modo de ejercicio.  
*ése* = el correr de Juan todos los días.
- (2) Juan corre todos los días de la semana.  
*Lo hace* por razones de salud.  
*lo hace* = acto de correr todos los días de la semana.

El orden temporal relativo se ve afectado por el tiempo y el aspecto verbales, así como por modificadores adverbiales; también depende de cambios en el estado atencional. En el discurso siguiente (de Webber 1987) el evento descrito en la preferencia 1, E1, ocurre en algún punto en el pasado, la preferencia 2 describe un evento E2 que ocurre antes que E1, el evento E3 de la preferencia 3 ocurre tras E2 y antes que E1 y está situada en la localización descrita en la preferencia 2, y el evento E4 de la preferencia 4 ocurre después de E1.

- (1) Juan fue a casa de María.
- (2) En el camino, se detuvo en una floristería a por rosas.
- (3) Cogió 5 rojas y 3 blancas.
- (4) Lamentablemente no consiguieron animarla.

Dos de las propuestas principales de una teoría de referencia a eventos, la de Webber (1987) y la de Hinrichs (1986) y Partee (1984), proporcionan una explicación del orden relativo de los eventos. Webber también proporciona una explicación del reconocimiento del orden de los eventos basada en el foco temporal y el movimiento de foco, conceptos análogos al del foco inmediato (Sidner 1983, Grosz et al. 1983) para la anáfora definida.

### **Relaciones de modificación en los sintagmas nominales.**

Las relaciones de modificación entre partes de un sintagma nominal complejo plantean dos problemas para la interpretación o generación de referencia. El primero es relativo a la selección de una única descripción subyacente (intendida) de las relaciones

estructurales y funcionales entre las entidades descritas por el sintagma nominal complejo; esta descripción única debe seleccionarse entre un número de tales descripciones posibles. Por ejemplo, la descripción intendida del sintagma *the pump dispenser* descansa en descubrir la relación funcional entre los dos objetos: la bomba y la máquina. El segundo problema para interpretar las relaciones de modificación es identificar la entidad referida (usando, claro está, alguna técnica para determinar las relaciones posibles). Los problemas de identificación pueden ser bastante difíciles para sintagmas nominales complejos con modificadores preposicionales usados en un contexto que incluya varios posibles referentes para cada parte simple del sintagma nominal. El problema se ilustra con el sintagma *the cat in the hat* usado en un contexto donde hay dos gatos, dos sombreros, pero sólo un gato en un sombrero (uno de *los* sombreros). Aunque el referente esté aquí claro, ninguna búsqueda simple de un único gato o bien un único sombrero en el contexto puede servir como base para la búsqueda del referente.

La investigación de los sintagmas nominales complejos ha explorado marcos para usar información lingüística con objeto de predecir la posible estructura de sintagmas nominales complejos (comparar con Isabelle 1984). Las aproximaciones computacionales a estos problemas se han ocupado de los problemas conjuntos de representar apropiadamente el conocimiento del dominio necesario para determinar la interpretación intendida y de definir procesos para la búsqueda de este conocimiento.

En contraste con éstas, las aproximaciones recientes (Mellish 1982, Haddock 1987) a la identificación del referente (que han considerado sintagmas preposicionales adjuntos más que nominales complejos) se han concentrado en el concepto de evaluación incremental de referencia usando técnicas de satisfacción de restricciones. Estas técnicas asumen que el núcleo nominal, el sintagma nominal incrustado en el sintagma preposicional y la relación especificada por la preposición forman un conjunto de restricciones sobre la elección de referentes candidatos para el sintagma nominal entero; estas restricciones pueden aplicarse incrementalmente a medida que se construye el sintagma para buscar el referente intenido. Estas aproximaciones asumen alguna representación del estado atencional que delimita un pequeño conjunto de referentes posibles y sirve para limitar la búsqueda emprendida por la satisfacción incremental de restricciones. Sólo muy recientemente

(Pollack and Pereira 1988) se ha explorado una teoría que incorpora los aspectos de la evaluación incremental con la búsqueda de métodos para nominales complejos discutidos previamente.

### **Elipsis discursiva**

La elipsis discursiva es la omisión (o *elisión*) en una preferencia de un sintagma requerido sintácticamente cuando el contenido necesario para determinar la interpretación de la preferencia elidida puede recuperarse de una preferencia previa. El tratamiento adecuado de la elipsis discursiva requiere que se considere el discurso (en vez de la oración) como unidad primaria de comunicación. Se han investigado dos tipos de elipsis discursiva. En el primer tipo el material elidido puede recuperarse directamente de la representación del significado de la preferencia previa. En el segundo la estructura intencional proporciona la fuente del material elidido; en este caso el contenido elidido puede no aparecer directamente en ninguna preferencia anterior.

Las aproximaciones del primer tipo (Hendrix et al. 1978, Webber 1983) han considerado la elipsis tanto nominal como verbal. Han considerado la elipsis como un fenómeno similar a la anáfora discursiva porque la elipsis también se interpreta estableciendo sintagmas mencionados previamente y sus representaciones discursivas como la fuente para reconstruir el material. De modo diferente al caso de la anáfora discursiva, los sintagmas elididos no deben referirse al mismo acto o entidad que el sintagma mencionado previamente, como se muestra en el ejemplo siguiente:

(1) Fred besó a su madre.

John también 0.

0 = besó a su propia madre o, alternativamente, besó a la madre de Fred.

En vez de eso la representación discursiva del sintagma mencionado previamente proporciona el material para construir la representación del sintagma elidido<sup>8</sup>.

Las aproximaciones del segundo tipo han utilizado la estructura intencional como la fuente para el material elidido.

---

<sup>8</sup> Esta concepción contrasta con la de Halliday y Hasan (1976), que ven la elipsis como una forma de sustitución textual.

Las preferencias elididas se ven como contribuciones al propósito general de un discurso o un segmento discursivo. Mientras el entorno discursivo contenga información suficiente no se necesita utilizar una preferencia completa; un sintagma o una parte de una oración contribuirá lo mismo. Fragmentos elípticos como (1) (de Allen 1979) y (2) (de Carberry 1985) abajo pueden verse como contribuciones directas al propósito discursivo; no se necesita reconstrucción alguna de ninguna oración completa en la que el fragmento esté incrustado.

(1) (dicho a la persona de la cabina de información en una estación del ferrocarril) ¿El tren a Windsor?

(2) Quiero obtener un título. Licenciado en Informática. Ningún curso por la noche.

Allen (1979) demostró que tales sintagmas pueden asociarse directamente con un acto de habla que a su vez identifica una parte del plan del hablante para actuar en un dominio de interés. Los modelos de reconocimiento de planes de Allen y Perrault (1980), Carberry (1985) y Litman (1985) recalcaron la posición de Allen y vincularon el fragmento con actos de habla y con los planes de los hablantes (esto es, planes en un dominio de acción). Concibiendo la acción como la raíz de la explicación de por qué ocurre el discurso, estos investigadores ilustraron que algunas formas de elipsis pueden verse en términos de la estructura intencional del discurso en lugar de en términos de un conjunto de relaciones con otras preferencias.

## 11.5 Reconocimiento de planes

Las explicaciones contemporáneas de reconocimiento de planes en la comprensión del discurso se derivan en gran medida de dos grandes teorías del filósofo Grice: su teoría del *significado no natural*<sup>9</sup> (Grice 1957, 1968) y su teoría de la *implicatura* (Grice 1975, 1978). En la teoría del significado no natural de Grice lo que los hablantes comunican cuando utilizan el lenguaje depende crucialmente de sus intenciones. La teoría de la implicatura descansa sobre la observación de que mucho de lo que se comunica intencionalmente en el uso del lenguaje no se expresa explícitamente; la misma teoría muestra cómo los hablantes pueden comunicar más de lo que expresan

---

<sup>9</sup> Los significados no naturales incluyen el tipo de significados transmitidos en el lenguaje y deben distinguirse de los significados naturales, por ejemplo, que el humo “significa” fuego.

explícitamente en sus preferencias. En ambas teorías se ve a los hablantes como productores de preferencias con la intención de que sus oyentes reconozcan las intenciones subyacentes a esas preferencias. La colección de intenciones subyacentes a cualquier preferencia incluirá pues, pero no estará limitada a, una intención de que al menos una parte de tal colección sea reconocida.

La comprensión del lenguaje requiere pues determinar las intenciones que tienen los hablantes: hacerse una idea de qué planes persiguen, en parte realizando una preferencia. El reconocimiento de planes en la conversación es una tarea viable precisamente porque el hablante tiene la intención de que el oyente la realice. El hablante no puede conseguir el efecto perseguido a menos que el oyente reconozca el plan del hablante; por ello el hablante incluirá en la preferencia lo que crea ser información suficiente para que el reconocimiento del plan sea posible para el oyente. Esta característica del reconocimiento de planes en la conversación lo distingue del problema de determinar los planes de un agente simplemente mediante la observación de sus acciones sin interaccionar con él. Este último problema, llamado “reconocimiento del ojo de la cerradura” por su semejanza con el problema de mirar a un agente a través del ojo de la cerradura e inferir su siguiente acción, es en general mucho más difícil<sup>10</sup>.

Al construir modelos detallados del proceso de reconocimiento de planes en el discurso, los investigadores no sólo han recurrido a la obra de Grice sino también a las ideas de Austin (1962) y de Searle (1969, 1975) de que el lenguaje se utiliza no sólo para “decir” cosas sino también para “hacer” cosas. En otras palabras, el comportamiento comunicativo debe verse primariamente como acción intencional. Se sigue de ello que en gran medida se pueden aplicar al análisis de la acción lingüística los mismos instrumentos que se usan para analizar la acción no lingüística—una idea desarrollada originalmente en gran medida por el trabajo de Cohen y Perrault (1979), Allen (1983) y Perrault y Allen (1980). En particular, los modelos de IA de reconocimiento de planes en el discurso han hecho uso de técnicas de representación y razonamiento sobre la acción originalmente desarrolladas por aquéllos interesados en el

---

<sup>10</sup> Para ejemplos de sistemas para el reconocimiento del ojo de la cerradura, véase Fischer et al. 1985, Genesereth 1979, McCue y Lesser 1983, y Schmidt et al. 1978.

problema de la formación automática de planes<sup>11</sup>. No obstante, antes de pasar a la discusión de estos modelos, merece la pena señalar el rango de fenómenos discursivos para los que es importante el reconocimiento de planes.

Considérese una persona que se acerca a la persona de la ventanilla de información de la estación de trenes y le dice, “¿Sabe usted cuándo sale el próximo tren a Detroit?” con la intención de que se le diga la hora de salida del siguiente tren a Detroit<sup>12</sup>. El plan comunicativo del hablante es hacer una petición—que se le diga cuando sale el siguiente tren a Detroit—y hacer tal petición profiriendo su pregunta. Si la persona de la ventanilla reconoce este plan, puede responder cooperativamente realizando la acción solicitada (asumiendo, claro está, que es capaz de hacerlo y no tiene ninguna razón para preferir no hacerlo). Es decir, puede contestar diciendo “A las 12:30” en lugar de simplemente “Sí”.

Normalmente los actos comunicativos se realizan como parte de planes mayores, y a menudo es necesario para un oyente inferir también esos planes mayores. En este ejemplo si la persona de información reconoce que el plan comunicativo del hablante es probablemente parte de un plan para ir a Detroit, puede proporcionarle a éste información adicional que le facilite su objetivo. Puede decirle, por ejemplo, de qué vía sale el siguiente tren a Detroit o puede criticar su plan, diciéndole que aunque el siguiente tren a Detroit sale en 15 minutos, es un tren local, y que el expreso que sale en 45 minutos le llevará antes a su destino.

El reconocimiento de planes es pues importante para responder adecuadamente a las preferencias del hablante en el discurso interactivo<sup>13</sup>. Es también importante para comprender la coherencia de preferencias de cláusulas múltiples. Considérese

---

<sup>11</sup> La obra de Austin y de Searle ha sido también terriblemente influyente fuera de la IA, inspirando todo un campo de investigación llamado *teoría de los actos de habla*, con investigadores en cada una de las principales disciplinas de la ciencia cognitiva. Levinson (1983, capítulo 5) proporciona un buen resumen de la obra en teoría de los actos de habla. Véanse también los artículos de Cohen, Morgan y Pollack (1990).

<sup>12</sup> Los ejemplos de la “estación de tren” han sido muy discutidos en la literatura. Fueron inspirados por un conjunto de diálogos grabados en el puesto de información de la estación de Toronto (Horrigan 1977). El ejemplo discutido aquí fue analizado en primer lugar por Allen (1983).

<sup>13</sup> La generación de respuestas adecuadas puede depender también de distinguir entre respuestas entendidas y no entendidas (aunque útiles) (Sidner 1983, 1985).

una persona distinta que le dice a la de la ventanilla de información “Voy a Detroit. ¿Dónde está la vía 7?” Es reconociendo el plan que pueda tener el hablante como puede determinar por qué es coherente la preferencia del hablante; efectivamente, el reconocimiento del plan es lo que le permite comprender completamente la preferencia. Es esencial para poder entender la coherencia de preferencias de múltiples cláusulas en el discurso tanto interactivo como no interactivo. De hecho varios estudios de reconocimiento de planes se han centrado en la comprensión de relatos, una forma de discurso esencialmente no interactivo (Bruce 1981, Wilensky 1983).

### La técnica básica

Como se señaló en la subsección anterior, la idea de que el uso del lenguaje debería ser visto como acción intencional condujo a modelos del discurso que hacen uso de técnicas ya desarrolladas para el razonamiento sobre las acciones (no lingüísticas). En particular la representación de planes y acciones utilizada en la mayoría de sistemas de reconocimiento de planes es una extensión directa de la representación desarrollada por primera vez en el sistema STRIPS (Fikes and Nilsson 1971) y luego extendida en el sistema NOAH (Sacerdoti 1977). Dado que la representación STRIPS tiene ciertas limitaciones de expresividad, ha habido desde mediados de los 80 un resurgimiento del interés en el proceso de formación de planes, y se han elaborado diversos esquemas de representación alternativos<sup>14</sup>. Sin embargo, las técnicas básicas de reconocimiento de planes se han desarrollado para representaciones de acciones del estilo de STRIPS.

En estas representaciones los estados del mundo se modelan como conjuntos de proposiciones, y las acciones se modelan como *operadores* que hacen corresponder un estado del mundo con otro. Cada operador tiene un *encabezado*, que da su nombre a la acción representada, una *lista de precondiciones*, que describe las proposiciones que han de ser verdaderas para que se pueda realizar la acción, y una *lista de efectos*, que describe lo que resulta verdadero por la realización de la acción. Un *plan* se analiza como una secuencia de acciones y estados del mundo. Concretamente  $[\alpha_1, S_1, \dots, S_{n-1}, \alpha_n]$  es un plan para transformar un estado  $S_0$  en un estado  $S_n$  bajo el supuesto de que

---

<sup>14</sup> Georgeff (1987) proporciona un buen repaso a estos desarrollos recientes; Georgeff y Lansky (1986a) han editado una colección de artículos importantes.

- todas las precondiciones de  $\alpha_1$  son verdaderas en  $S_0$ ;
- todos los efectos de  $\alpha_n$  son verdaderas en  $S_n$ ;
- en cada estado intermedio  $S_i$  todos los efectos de  $\alpha_i$  y todas las precondiciones de  $\alpha_{i+1}$  son verdaderas.

La formación de planes es el proceso de encontrar un plan que transforme el estado presente del mundo en un estado en el que algún conjunto de objetivos sea cierto. En el caso más simple un plan puede proyectarse como un proceso de *grafo-búsqueda*, en el que los nodos del grafo representan acciones o estados del mundo, y las conexiones lícitas entre los nodos se determinan por las relaciones entre estados y acciones especificadas en el conjunto de operadores (ver capítulo 7 de Nilsson 1980).

Esta concepción simple puede refinarse permitiendo que cada operador incluya un *cuerpo* así como una lista de precondiciones y acciones. El cuerpo de un operador puede ser una lista de subacciones cuya realización constituye la realización de la acción nombrada por el encabezado o bien una lista de subobjetivos cuya consecución constituye la realización de la acción del encabezado. Cuando los operadores incluyen cuerpos, el proceso de formación de planes puede ser *jerárquico*: puede buscarse primero una secuencia de acciones y estados intermedios que son relativamente abstractos y después elaborarla en planes más y más detallados. La formación jerárquica de planes es a menudo mucho más eficiente que la formación lineal de planes (Sacerdoti 1977, Stefik 1981, Tate 1984, Wilkins 1984)<sup>15</sup>.

En reconocimiento de planes un agente observa a otro que realiza alguna acción o secuencia de acciones e intenta determinar qué plan está siguiendo el último. En el discurso las acciones observadas son las preferencias de un hablante<sup>16</sup>; así  $H$  (de hablante) denotará al agente cuyo plan está siendo inferido y  $O$  (de oyente) denota al agente que está infiriendo el plan de  $H$ . Supongamos que  $H$  observa a  $O$  realizando una acción  $\alpha$  en un estado  $S_1$  en el que todas las precondiciones de  $\alpha$  son verdaderas. Entonces  $O$  puede concluir que el objetivo de  $H$  al realizar la acción  $\alpha$  es conseguir uno o varios de los efectos de  $\alpha$ . De hecho

---

<sup>15</sup> Rosenschein (1981) proporciona un análisis lógico de planes jerárquicos; Wilkins (1985) clarifica diversas distinciones importantes que deben hacerse entre procesos de planificación jerárquicos.

<sup>16</sup> Naturalmente aquí, como a lo largo de este capítulo, por “hablante” nos referimos al agente que produce lenguaje, sea ese lenguaje oral o escrito.

el razonamiento de  $O$  puede ir más allá. Si algún efecto de  $\alpha$ , digamos  $e$ , es una precondition de otra acción  $\beta$ ,  $O$  puede concluir que  $H$  tiene la intención de realizar  $\beta$ . El proceso puede repetirse:  $O$  puede concluir que  $H$  tiene la intención de conseguir algún efecto de  $\beta$ , etcétera. Generalmente este proceso de razonamiento termina cuando  $O$  determina que  $H$  tiene la intención de conseguir algún objetivo de dominio típico.

Una de las primeras y más influyentes formulaciones del reconocimiento de planes fue la de Allen (1979, 1983). Allen proporcionó un conjunto de reglas de reconocimiento de planes junto con una estrategia heurística para controlar su aplicación. Una regla típica de reconocimiento de planes es la Regla Acción-Efecto.

$$\text{Bel}(O, \text{Int}(H, \alpha)) \rightarrow \text{Bel}(O, \text{Int}(H, e)) \quad \text{si } e \text{ es un efecto de } \alpha.$$

Esta regla puede glosarse como “si el oyente (o de modo más general el agente que infiere) cree que el hablante (o de modo más general el que actúa) tiene la intención de realizar alguna acción  $\alpha$ , entonces el oyente puede decidir que el hablante tiene la intención de hacer que cierta proposición  $e$  sea verdadera, si  $e$  es un efecto de la realización de  $\alpha$  (en el estado del mundo en que se realice  $\alpha$ ).”<sup>17</sup>

Para que el proceso de reconocimiento de planes comience es necesario un supuesto de intencionalidad. Esto es, dado que  $O$  observa que  $H$  realiza una acción  $\alpha$ , la Regla Acción-Efecto sólo se aplica si  $O$  concluye que  $S$  realizó  $\alpha$  intencionalmente. El supuesto de intencionalidad no es problemático en casos en los que el reconocimiento de planes se realiza en conversación: si  $H$  dice algo a  $O$ , es bastante razonable que  $O$  suponga que la preferencia de  $H$  era intencional.

Otras reglas vinculan proposiciones a acciones de las que son preconditiones y subacciones a acciones que las contienen. Estas reglas, más la ya discutida, se usan para lo que se denomina *encadenamiento hacia delante*: razonamiento desde una acción observada hacia los efectos de esa acción, y desde estos efectos a acciones subsiguientes que se han hecho posibles, etcétera. Los sistemas de reconocimiento de planes también incluyen

---

<sup>17</sup> En realidad una regla como la de Acción-Efecto se utiliza para inferir no sólo que  $H$  tiene la intención de hacer que  $e$  sea verdadera, sino que  $H$  tiene la intención de conseguir  $e$  haciendo  $\alpha$ —esto es, que el plan de  $H$  incluye a la subsecuencia  $[\alpha, e]$ .

normalmente reglas para el *encadenamiento hacia atrás*. Un agente que infiere puede razonar también acerca de los objetivos que el que actúa probablemente tiene y desde ahí hacia atrás acerca de que acciones que satisfarían esos objetivos, cuáles son las precondiciones de esas acciones, etcétera. La Regla Efecto-Acción de Allen es un ejemplo de una regla de inferencia de planes para utilizar en el encadenamiento hacia atrás:

$$\text{Bel}(O, \text{Int}(H, e)) \rightarrow \text{Bel}(O, \text{Int}(H, \alpha)) \quad \text{si } e \text{ es un efecto de } \alpha.$$

Además de las reglas simples de encadenamiento hacia delante y hacia atrás, muchos sistemas de reconocimiento de planes tienen reglas adicionales que se aplican sólo a acciones de búsqueda de información. Por ejemplo, muchos sistemas relacionan la acción de “determinar si  $p$ ”, para alguna proposición  $p$ , con la acción de “conseguir que  $p$ .” Estas reglas, junto con reglas para manipular lo que se denomina inferencia de planes anidados, sin embargo, pueden verse como casos particulares de reglas más simples (Kautz 1985).

Es importante señalar que las reglas de inferencia de planes no deben interpretarse como implicaciones lógicas sino más bien como descripciones de inferencias “probables” pero no deductivas que un agente puede hacer al realizar reconocimiento de planes. Kautz (1990) presenta una formalización precisa del proceso de utilización de estos principios de inferencia probable para realizar reconocimiento de planes, proporcionando teorías de modelos, de demostración y algorítmicas fundadas en la teoría de la circunscripción de McCarthy (McCarthy 1980, 1984).

La aplicación de las reglas de inferencia de planes se controla normalmente por un conjunto de heurísticas, que se designan tanto para hacer que el proceso de reconocimiento sea más eficiente como para asegurar que los planes más verosímiles se encuentren antes que los menos verosímiles. Algunas de las heurísticas de control se basan en nociones de sentido común sobre la naturaleza de los planes y se aplican igualmente bien al proceso de inferir los planes del interlocutor conversacional de uno y a inferir los planes de un agente que uno está meramente observando. Por ejemplo, una heurística tal dirige un sistema de reconocimiento de planes contra la tendencia de considerar como candidatos a planes que contengan acciones cuyos efectos sean verdaderos en el momento en que se vaya a realizar la acción. Esta heurística es razonable tanto cuando el reconocimiento de planes se produce en el discurso como cuando no: en general los planes no contienen acciones superfluas.

Por otra parte, un número importante de heurísticas de control sólo se justifica en reconocimiento de planes en conversación al fundarse en la noción griceana de *reconocimiento intencido*: que el hablante tiene la intención de que sus oyentes reconozcan al menos un subconjunto de sus intenciones. Esta idea motiva heurísticas de reconocimiento de planes tales como la *heurística de bifurcación*, que afirma que la probabilidad de que un plan candidato sea el que un hablante tiene realmente es inversamente proporcional al número de alternativas a las que puede extenderse (Allen 1983), y la del *supuesto de rama única*, que afirma que si un hablante cree que en cierta fase del discurso se podría inferir más de un plan, es su responsabilidad dar a conocer aquél que tiene—el oyente sólo debe inferir hasta el punto de una ruptura potencial (Sidner 1985). Tales heurísticas median contra inferencias difíciles sobre el supuesto de que los hablantes en general intentan producir preferencias que facilitan un reconocimiento de planes relativamente simple.

Sólo como un ejemplo de la aplicación de las técnicas básicas de reconocimiento de planes en el discurso, considérese a un viajero que dice, “¿Cuándo sale el próximo tren a Detroit?” De acuerdo con Allen, podemos codificar esta acción preferencial como

SOLICITAR(*H*, *O*, INFORM-REF(*H*, *O*, tiempo1)).

Es decir, *H* ha solicitado a *O* que le informe de la referencia de cierta constante tiempo1, que denota la hora de partida del tren siguiente a Detroit<sup>18</sup>. Aplicando la Regla Acción-Efecto al operador SOLICITAR

Encabezado: SOLICITAR(Agente1, Agente2, Acción)

Efecto: QUERER(Agente2, HACER(Agente2, Acción))

*O* puede concluir que *H* puede querer que *O* quiera informar a *H* de la hora de salida del tren a Detroit. El razonamiento ulterior permite que *O* finalmente infiera el plan entero que se muestra en la figura 1.1.<sup>19</sup>

---

<sup>18</sup> La codificación de los actos de habla de Allen se deriva en gran medida de la formalización de Cohen de la teoría de actos de habla, que utiliza nociones de planificación IA (Cohen y Perrault 1979).

<sup>19</sup> La mayoría de los operadores que se muestran en la figura 11.1 deberían ser autoexplicativos. SABER-REF(agente, valor) debe entenderse como “el agente sabe la referencia de (o el valor de) la constante.” Los argumentos

EN( $H$ , detroit)  
     ↑ acción-efecto  
 VIAJAR-A( $H$ , detroit, tren1)  
     ↑precondición-acción  
 A-BORDO( $H$ , tren1)  
     ↑ acción-efecto  
 EMBARCAR( $H$ , tren1, detroit)  
     ↑precondición-acción  
 EN( $H$ , ubic1, tiempo1)  
     ↑ acción-efecto  
 IR-A( $H$ , ubic1, tiempo1)  
     ↑precondición-acción  
 SABER-REF( $H$ , tiempo1)  
     ↑ acción-efecto  
 INFORM-REF( $H$ ,  $O$ , tiempo1)  
     ↑precondición-acción  
 QUERER( $O$ , INFORM-REF( $H$ ,  $O$ , tiempo1))  
     ↑ acción-efecto  
 SOLICITAR( $H$ ,  $O$ , INFORM-REF( $H$ ,  $O$ , tiempo1))  
  
 tiempo1=hora-salida (tren1)  
 ubic1=ubicación-salida (tren1)

Figura 11.1 Un plan de embarcar en un tren

Los arcos del grafo del plan se han etiquetado con reglas de inferencia de encadenamiento hacia delante, pero está claro que alguna parte del grafo en realidad pudo haberse encontrado utilizando encadenamiento hacia atrás. Para inferir este plan también se necesitan varios operadores adicionales. Uno es el operador EMBARCAR; la construcción de los demás se la dejamos al lector como ejercicio.

---

de los operadores que van en mayúsculas (excepto  $H$  y  $O$ ) son variables; los argumentos en minúsculas son constantes.

Encabezado: EMBARCAR(Agente, Tren, Estación)

Precondición: EN(Agente, ubicación-salida(Tren), hora-salida(Tren))

Efecto: A-BORDO(Agente, Tren)

### **Extensiones de la técnica básica**

La técnica básica para reconocimiento de planes tiene varias limitaciones. Quizá la más notable sea el hecho de que se limita a casos en los que el agente que infiere hace una sola observación. Cuando se aplica al discurso, esto significa que la técnica básica es útil para razonamiento sólo sobre preferencias de una única cláusula. Pero una de las características definitorias de los discursos es que se prolongan en el tiempo. Los hablantes transmiten información usando varias cláusulas y en el discurso interactivo usando varias preferencias (o “turnos”).

Las *técnicas de reconocimiento incremental* se han desarrollado para extender la técnica básica de reconocimiento de planes para tratar con secuencias de preferencias (Carberry 1988, Sidner 1983, 1985). En el reconocimiento incremental el sistema que desempeña el papel de *O*, comienza usando la técnica básica para inferir tanto como sea posible de la primera preferencia (de inicio de discurso) de *H*. En ese momento puede no ser posible determinar cuál de los varios planes es el que sigue *H* y puede no ser posible determinar el plan de *H* a un nivel suficiente para responder apropiadamente. El resultado provisional puede ser entonces un conjunto de planes candidatos parciales. Una vez que se han computado éstos, el procesamiento se interrumpe. Al escuchar cada preferencia subsiguiente, el sistema trata de expandir todos los planes parciales ya construidos que pueda. El proceso de expansión hace uso otra vez de la técnica básica: las reglas de inferencia de planes se utilizan para relacionar un nodo de un subgrafo existente con la acción hecha al realizar la preferencia y de ahí razonar desde la acción a otras acciones. Puede probarse imposible expandir algunos de los subgrafos existentes: entonces éstos dejan de ser considerados, se eliminan.

El reconocimiento de planes incremental puede hacerse más eficiente tomando en cuenta el foco de atención del hablante. Como señalamos en la sección 11.2, al menos en los diálogos orientados a tarea, la estructura de la tarea tiene influencia sobre la estructura del discurso. Cuando un hablante habla sobre un plan de realizar cierta tarea, no es probable que “salte” en la presentación, hablando primero de cierta subtarea, luego sobre otra, para volver luego a la primera, y así

sucesivamente. Más bien, en cada punto del discurso, alguna parte del plan general del hablante será más prominente que otras es más probable que las preferencias del hablante tengan que ver con la parte prominente. Esta idea puede utilizarse para restringir el proceso de reconocimiento de planes: *O* puede preferir expandir aquellas partes de los subplanes candidatos que están en el foco que aquellas que no lo están (Carberry 1988, Litman y Allen 1989)<sup>20</sup>.

En procesamiento del discurso es a menudo necesario reconocer varios planes relacionados que un hablante está ejecutando. En particular, los hablantes a menudo interrumpen su discusión de un plan de dominio para ejecutar un plan de discurso, por ejemplo, el de pedir clarificación o corregir un malentendido anterior, con la ejecución de planes de dominio. En el diálogo siguiente, adaptado de uno de Litman y Allen 1989, el pasajero se ocupa de un plan de clarificar la información que ya ha recibido en medio de la realización de su plan de determinar la hora y ubicación de salida de un tren determinado:

Pasajero: ¿El tren a Montreal?

Empleado: Puerta siete.

Pasajero: ¿Dónde está?

Empleado: Por aquí abajo. La segunda a la izquierda.

Pasajero: ¿Y a qué hora sale?

Empleado: A las nueve.

Se pueden introducir operadores adicionales que representen planes de discurso. Litman ha propuesto el uso de una pila de planes parciales en reconocimiento de planes: las representaciones del discurso deben apilarse encima de las representaciones de planes de dominio de los que dependen (Litman y Allen 1989).

Cuando la conducta observada es el resultado de dos o más planes de dominio que interactúan surge una dificultad más. Wilensky (1983) ha propuesto tratar estos casos proporcionando al agente que infiere un conjunto de *metaplanes* que operan sobre otros planes de dominio para construir planes más complejos: *resolver-conflictos* es uno de esos metaplanes.<sup>21</sup> En

---

<sup>20</sup> También pueden usarse para restringir el reconocimiento incremental los indicadores lingüísticos de la estructura del discurso, como los *sintagmas de entrada*; véase la sección 11.3 para una discusión de tales indicadores.

<sup>21</sup> Algunos trabajos recientes en formación de planes también incorporan la idea de metaplanes como éstos; véase, por ejemplo, Georgeff and Lansky 1986b.

textos no-interactivos también hay un problema al razonar sobre interacciones entre los planes separados de agentes múltiples, por ejemplo, los personajes de una historia; este problema ha sido estudiado en detalle por Bruce y sus colaboradores (Bruce 1981).

Otro tipo de restricciones inherentes al modelo standard de reconocimiento de planes ha sido señalado por Pollack (1986, 1988): supone que el agente que infiere (*O*) y el que actúa (*H*) tienen creencias extremadamente similares sobre el dominio de la acción. Considérese otra vez la regla de inferencia de planes Acción-Efecto. La glosábamos como “si el *O* cree que el *H* tiene la intención de realizar una acción  $\alpha$ , entonces el *O* puede decidir que el *H* tiene la intención de que una proposición *e* sea verdadera, si *e* es un efecto de realizar  $\alpha$ .” Nótese que esta regla no señala precisamente quién es—*O* o *S*—quien cree que *e* es una precondition de  $\alpha$ . Si la tomamos como una creencia de *O*, no está claro que *O* infiera el plan de *H*; por otra parte, si la consideramos una creencia de *H*, no está claro como llega *O* a tener acceso directo a ella. En la práctica sólo hay un único conjunto de operadores que vincula las condiciones y las acciones en los modelos standard de reconocimiento de planes; la creencia en cuestión se considera tanto de *O* como de *H*.

En muchas situaciones el supuesto de que las creencias relevantes de *H* son idénticas a las de *O* resulta en un fracaso no sólo del proceso de reconocimiento de planes sino también del proceso comunicativo que el reconocimiento de planes pretende sustentar. En particular impide la generación por principio de respuestas apropiadas a preguntas que surgen de planes no válidos. Pollack (1986, 1988) ha propuesto un modelo de reconocimiento de planes en la conversación que distingue entre las creencias de *H* y las de *O*. El modelo descansa sobre un análisis de los planes como fenómeno mental: “tener un plan” se analiza como el hecho de tener una configuración particular de creencias e intenciones. El proceso de reconocimiento de planes puede beneficiarse de diversas técnicas de adscripción de creencias. Los juicios de que un plan es no válido se asocian con discrepancias concretas entre las creencias que *O* atribuye a *H*, cuando *O* cree que *H* tiene un plan, y las propias creencias de *O*.

Además la participación en el discurso puede requerir la habilidad de representar planes conjuntos de agentes múltiples. Grosz y Sidner (1990) han mostrado que estos planes no pueden definirse sólo en términos de los planes privados de los agentes individuales. Ellas han propuesto una representación que

extiende la definición de plan de Pollack a los planes de dos o más agentes cuando colaboran en una tarea y se comunican en el discurso.

Otra vía reciente en reconocimiento de planes en discurso ha sido el vincular más estrechamente las teorías de actos de habla con teorías fundamentales de la acción racional. Cohen y Levesque (1990) han mostrado que los actos ilocutivos, como solicitar o prometer, no necesitan ser tratados como primitivos en un modelo de reconocimiento de planes en discurso; en lugar de eso, la conducta propia de los participantes en el discurso puede analizarse en términos de un modelo de principios racionales de adopción de creencias e intenciones. De modo análogo, Perrault (1990) ha mostrado cómo una explicación de los actos de habla puede basarse en una teoría simple de adopción de creencias y observación de acciones. Como Cohen y Levesque, Perrault evita la necesidad de tratar los actos ilocutivos como primitivos en el modelo. Su teoría difiere en que hace uso de la lógica por defecto de Reiter (1980), que le permite presentar axiomas más independientes del contexto.

## Referencias.

Allen, J. F. 1979. *A Plan-Based Approach to Speech Act Recognition*. Technical report 131, University of Toronto, Toronto, Canada.

Allen, J. F. 1983. Recognizing intentions from natural language utterances. In M. Brady and R. C. Berwick, eds. *Computational Models of Discourse*. Cambridge, MA: MIT Press, pp. 107-166

Allen, J. F. and Perrault, C. R. 1980. Analyzing intentions in dialogue. *Artificial Intelligence* 15(3): 143-178.

Appelt, D. and Kronfeld, A. 1987. A computational model of referring. In *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, Milan, pp. 640-647.

Austin, J. L. 1962. *How to Do Things with Words*. 2nd ed. Cambridge, MA: Harvard University Press.

Bobrow, D., and the PARC Understander Group. 1977. Gus, a frame driven dialog system. *Artificial Intelligence* 8: 155-173.

Bruce, B. 1981. Plans and social action. In *Theoretical Issues in Reading Comprehension*. Hillsdale, NJ: Lawrence Erlbaum.

Butterworth, B. 1975. Hesitation and Semantic planning in speech. *Journal of Psycholinguistic Research* 4: 75-87.

Carberry, M. S. 1985. A pragmatics based approach to understanding intersentential ellipsis. In *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics*. Chicago, pp. 188-197.

Carberry, M. S. 1988. Pragmatic modeling: Toward a robust natural language interface. *Computational Intelligence* 3.

Chafe, W. L. 1979. The flow of thought and the flow of language. In T. Givon. ed. *Syntax and Semantics 12: Discourse and Syntax*. New York: Academic Press, pp. 159-182.

Chafe, W. L. 1980. The deployment of consciousness in the production of a narrative. In *The Pear Stories: Cognitive, Cultural and Linguistic Aspects of Narrative Production*. Norwood, NJ: Ablex, pp. 9-10.

Charniak, E. 1977. A framed painting: The representation of commonsense knowledge fragment. *Cognitive Science* 1(4).

Cohen, P. R., and Levesque, H. (1990), Rational interaction as the basis for communication. In P. R. Cohen, J. L. Morgan, and M. E. Pollack, eds. *Intentions in Communication*. Cambridge, MA: MIT Press.

Cohen, P. R., Morgan, J. L. and Pollack, M. E. eds. 1990 *Intentions in Communication*. Cambridge, MA: MIT Press.

Cohen, P. R. and Perrault, C. R. 1979. Elements of a plan-based theory of speech acts. *Cognitive Science* 3: 177-212.

Cohen, R. 1984. A computational theory of the function of the clue words in argument understanding. In *Proceedings of COLING84*. Stanford, CA.

Cohen, R. 1987. Analyzing the structure of argumentative discourse. *Computational Linguistics* 13(1-2).

Cullingford, R. E. 1977. *Script application: Computer understanding of newspaper stories*. Doctoral dissertation, Yale University, New Haven, CT.

Cullingford, R. E. 1981. Sam. In *Inside Computer Understanding: Five Programs Plus Miniatures*. Hillsdale, NJ: Erlbaum, pp. 75-119.

Dahl, D. A. 1987. Nominalizations in *PUNDIT*. In *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics*. Stanford, CA, pp. 131-139.

Donnellan, K. S. 1966. Reference and Definite Descriptions. *Philosophical Review* 75: 281-304. Reprinted in Steinberg and Jacobovits, eds. *Semantics*. Cambridge, Engl.: Cambridge University Press.

Fikes, R. E., and Nilsson, N. J. 1971. STRIPS: a new approach to the application of theorem proving to problem solving. *Artificial Intelligence* 2: 189-208.

Fillmore, C. J. 1975. *Santa Cruz Lectures on Deixis*. Technical report, Indiana Linguistics Club, Bloomington, IN.

Finin, T. 1980. *The Semantic Interpretation of Compound Nominals*. Ph. D. Thesis, University of Illinois, Urbana, IL.

Fischer, G., Lemke, A., and Schwab, T. 1985. Knowledge-based help systems. In *Proceedings of the CHI'85 Conference on Human Factors in Computing Systems*, ACM SIGCHI, San Francisco, pp. 161-167.

Genesereth, M. R. 1979. The role of plans in automated consultation. In *Proceedings of the 6th International Joint Conference on Artificial Intelligence*, Tokyo, pp. 311-319.

Georgeff, M. P. 1987. Planning. In J. Traub, ed. *Annual Review of Computer Science*. Palo Alto, CA: Annual Reviews.

Georgeff, M. P. and Lansky, A. L. 1986a. *Reasoning about Actions and Plans: Proceedings of the 1986 Workshop*. Los Altos, CA: Morgan Kaufmann.

Georgeff, M. P. and Lansky, A. L. 1986b. Procedural Knowledge. *Proceedings of the IEEE, Special Issue on Knowledge Representation*, pp. 1383-1398.

Grice, H. P. 1957. Meaning. *Philosophical Review* 67.

Grice, H. P. 1968. Utterer's meaning, sentence-meaning, and word-meaning. *Foundations of Language* 4: 1-18.

Grice, H. P. 1975. Logic and Conversation. In P. Cole and J. L. Morgan, eds. *Syntax and Semantics Vol. 3: Speech Acts*. New York: Academic Press, pp. 41-58.

Grice, H. P. 1978. Further notes on Logic and Conversation. In P. Cole, eds. *Syntax and Semantics Vol. 9: Pragmatics*. New York: Academic Press, pp. 113-128.

Grosz, B. 1974. The structure of task oriented dialogs. In *IEEE Symposium on Speech Recognition: Contributed Papers*, IEEE. Pittsburgh: Carnegie Mellon University, Computer Science Dept.

Grosz, B. J. 1977. The representation and use of focus in a system for understanding dialogs. In *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*, Cambridge, MA.

Grosz, B. J. 1978a. Discourse analysis. In D. Walker, ed. *Understanding Spoken Language*. Ch. IX. New York: Elsevier North-Holland, pp. 235-268.

Grosz, B. J. 1978b. Focusing in dialog. In *Proceedings of Workshop on Theoretical Issues in Natural Language Processing-2*, Urbana, IL.

Grosz, B. J. 1981. Focusing and description in natural language dialogues. In A. K. Joshi, I. Sag, and B. Webber, eds. *Elements of Discourse*. Cambridge, Engl.: Cambridge University Press.

Grosz, B. J., Joshi, A. K., and Weinstein, S. 1983. Providing a unified account of definite noun phrases in discourse. In *Proceedings of the 9th International Joint Conference on Artificial Intelligence*, Cambridge, MA.

Grosz, B. J. and Sidner, C. L. 1986. Attention, intentions and the structure of discourse, *Computational Linguistics* 12(3).

Grosz, B. J. and Sidner C. L. 1990. Plans for Discourse. In P. R. Cohen, J. Morgan and M. E. Pollack, eds. *Intentions in Communication*. Cambridge, MA: MIT Press.

Guenther, F., Lehmann, H., and Schoenfeld, W. 1986. A theory for representation of knowledge. *IBM Journal of Research and Development* 30(1): 39-56.

Haddock, N. J. 1987. Incremental interpretation and combinatory categorial grammar. In N. Haddock, E. Klein, and G. Morrill, eds. *Edinburgh Working Papers in Cognitive Science, Vol. 1: Categorical Grammar, Unification Grammar and Parsing*. Edinburgh, Scotland, pp. 71-84.

Halliday, M., and Hasan, R. 1976. *Cohesion in English*. English Language Series, Title No. 9. London: Longman's.

Heim, I. 1982, The semantics of definite and indefinite noun phrases. Ph. D. dissertation, University of Massachusetts, Amherst, MA.

Hendrix, G., Sacerdoti, E., Sagalowicz, D., and Slocum, J. 1978. Developing a natural language interface to complex data. *ACM Trans. on Database Sys.* 3(2).

Hinrichs, E. 1986. Temporal anaphora in discourses of English. English Language Series, Title No. 9. London: Longman's.

Hirschberg, J. B. and Pierrehumbert, J. 1986. The intonational structuring of discourse. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, New York, pp. 136-144.

Hirschberg, J. B. and Litman, D. J. 1987. Now let's talk about now: Identifying cue phrases intonationally. In *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics*, Stanford, CA, pp. 163-171.

Hirschberg, J. B., Litman, D. J., Pierrehumbert, J., and Ward, G. L. 1987. Intonation and the intonational structure of discourse. In *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, Milan, Italy, pp. 636-639.

Hobbs, J. 1979. Coherence and co-reference. *Cognitive Science* 1: 67-82.

Hobbs, J. 1983. Why is discourse coherent? In F. Neubauer, ed., *Coherence in Natural Language Texts*. Hamburg: H. Buske Verlag.

Hobbs, J. R. 1985. *On the Coherence and Structure of Discourse*. Technical Report CSLI-85-37, CSLI, Stanford, CA.

Hobbs, J. and Martin P. 1987. Local pragmatics. In *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, Milan, Italy.

Horrigan, M. K. 1977. *Modelling simple dialogues*. Technical report TR 108, Dept. of Computer Science, University of Toronto, Toronto, Ontario.

Isabelle, P. 1984. Another look at nominal compounds. In *Proceedings of the 10th International Conference on Computational Linguistics*, Stanford, CA, pp. 509-516.

Joshi, A. and Weinstein, S. 1981. Control of inference: Role of some aspects of discourse structure-centering. In *Proceedings of International Joint Conference on Artificial Intelligence*, Vancouver, BC, pp. 385-387.

Kamp, H. 1981. A theory of truth and semantic representation. In J. Groenendijk, T. V. M. Janssen, and M. Stokhof, eds. *Formal Methods in the Study of Language*. Mathematical Centre Tract 135, Amsterdam, The Netherlands, pp. 277-322.

Kautz, H. A. 1985. *Toward a Theory of Plan Recognition*. Technical report TR162, University of Rochester, Rochester, NY.

Kautz, H. A. 1990. A circumscriptive theory of plan recognition. In P. R. Cohen, J. L. Morgan, and M. E. Pollack, eds. *Intentions in Communication*. Cambridge, MA: MIT Press.

Kendon, A. 1972. Some relationships between body motion and speech. In *Studies in Dyadic Communication*. New York, Pergamon.

Kronfeld, A. 1986. Donnellan's distinction and a computational model of reference. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, New York.

Lasnik, H. 1976. Remarks on co-reference. *Linguistic Analysis* 2(1): 1-22.

Lehnert, W. 1977. A conceptual theory of question-answering. In *Proceedings of the 5th International Joint Conference on Artificial Intelligence*, Palo Alto, CA: Morgan Kaufmann.

Lehnert, W. 1981. Plot units and narrative summarization. *Cognitive Science* 5(4): 291-331.

Levinson, S. C. 1983. *Pragmatics*. Cambridge, Engl.: Cambridge University Press.

Levy, D. M. 1979. Communicative goals and strategies: Between discourse and syntax. In T. Givon, ed. *Discourse and Syntax*. New York: Academic Press, pp. 337-354.

Linde, C. 1979. Focus of attention and the choice of pronouns in discourse. In T. Givon, ed. *Syntax and Semantics*. New York: Academic Press, pp. 337-354.

Linde, C. and Goguen, J. 1978. Structure of planning discourse. *J. Social Biol. Struct.* 1: 219-251.

Litman, D. 1985. *Plan Recognition and Discourse Analysis: An Integrated Approach for Understanding Dialogues*. Technical report TR170, University of Rochester, Rochester, NY.

Litman, D. J. and Allen, J. F. 1987. A plan recognition model for subdialogues in conversations, *Cognitive Science* 11, 163-200.

Litman, D. J. and Allen, J. F. 1990. Discourse processing and commonsense plans. In P. R. Cohen, J. L. Morgan, and M. E. Pollack, eds. *Intentions in Communication*. Cambridge, MA: MIT Press.

Mann, W. D., Moore, M. A., Levin, J. A., and Carslile, J. H. 1975. *Observation Methods for Human Dialogue*. Technical report, Information Services Institute, Marina del Rey, CA.

Mann, W. D. and Thompson, S. A. 1986. Relational propositions in discourse. *Discourse Processes* 9(1): 57-90.

Marslen-Wilson, W., Levy, E., and Tyler, L. K. 1982. Producing interpretable discourse: The establishment and maintenance of reference. In *Speech, Place, and Action*. Chichester, Engl.: Wiley.

McCarthy, J. 1980. Circumscription—a form of non-monotonic reasoning. *Artificial Intelligence* 13: 27-39.

McCarthy, J. 1984. Applications of circumscription to formalizing common sense knowledge. In *Proceedings of the AAAI Workshop on Nonmonotonic Reasoning*, New Paltz, NY, pp. 295-324.

McCue, D. and Lesser, V. 1983. *Focusing and Constraint Management in Intelligent Interface Design*. Technical report COINS 83-36, University of Massachusetts, Amherst, MA.

McDonald, D. D. 1983. Natural language generation as a computation problem. In *Computational Models of Discourse*, Cambridge, MA: MIT Press.

McKeown, K. 1985. *Text Generation*. New York: Cambridge University Press.

McNeill, D. 1979. *The Conceptual Basis of Language*. Hillsdale, NJ: Erlbaum.

McNeill, D. and Levy, E. 1982. Conceptual representations in language activity and gesture. In *Speech, Place, and Action*. Chichester, Engl.: Wiley.

Mellish, C. S., 1982. *Incremental Evaluation: An Approach to the Semantic Interpretation of Noun Phrases*. Technical report, University of Sussex Cognitive Studies Programme, Sussex, Engl.

Nilsson, N. J. 1980. *Principles of Artificial Intelligence*. Palo Alto, CA: Tioga.

Partee, B. H. 1984. Nominal and temporal anaphora. *Linguistics and Philosophy* 7: 243-286.

Perrault, R. C. 1990. An application of default logic to speech act theory. In P. R. Cohen, J. L. Morgan, and M. E. Pollack, eds. *Intentions in Communication*. Cambridge, MA: MIT Press.

Perrault, R. C. and Allen, J. F. 1980. A plan-based analysis of indirect speech acts, *American Journal of Computational Linguistics* 6: 167-182.

Polanyi, L. 1985. A theory of discourse structure and discourse coherence. In W. Eilfort, P. Kroerber, and K. Peterson, eds. *Proceedings of the 21st Regional Meeting of the Chicago Linguistics Society*. Chicago: University of Chicago Press, pp. 306-322.

Polanyi, L. 1986. *The Linguistic Structure of Discourse: Towards a Formal Theory of Discourse Structure*. Technical report 6489, Bolt, Berenek and Newman, Cambridge, MA.

Polanyi, L. and Scha, R. 1984. A syntactic approach to discourse semantics. In *Proceedings of the 10th International Conference on Computational Linguistics*, Stanford, CA, pp. 413-419.

Pollack, M. E. 1986. A model of plan inference that distinguishes between the beliefs of actors and observers. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, New York, pp. 207-214.

Pollack, M. E. 1990. Plans as Complex Mental Attitudes. In P. R. Cohen, J. Morgan and M. E. Pollack, eds. *Intentions in Communication*. Cambridge, MA: MIT Press, pp. 77-103.

Pollack, M. E. and Pereira F. C. N. 1988. An integrated framework for semantic and pragmatic interpretation. In *Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics*, Buffalo, NY.

Reichman-Adar, R. (1984), Extended person-machine interface. *Artificial Intelligence* 22: 157-218.

Reinhart, T. 1983. Co-reference and bound anaphora: A restatement of the anaphora question. *Linguistics and Philosophy* 6: 47-88.

Reiter, R. 1980. A logic for default reasoning. *Artificial Intelligence* 13: 81-132.

Robinson, A. 1981. Determining verb phrase referents in dialogs. *Computational Linguistics* 7(1).

Rosenschein, S. J. 1981. Plan synthesis: a logical perspective. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence*. IJCAI, Vancouver, BC, pp. 331-337.

Rumelhart, D. E. 1975. Notes on a schema for stories. In D. G. Bobrow and A. Collins, eds. *Representation and Understanding*. New York: Academic Press.

Sacerdoti, E. D. 1977. *A Structure for Plans and Behavior*. New York: American Elsevier.

Schiffirin, D. 1982. *Discourse Markers: Semantic Resource for the Construction of Conversation*. Ph. D. Thesis, University of Pennsylvania, Philadelphia, PA.

Schmidt, C. F., Sridharan, N. S. and Goodson, J. L. 1978. The plan recognition problem: An intersection of artificial intelligence and psychology. *Artificial Intelligence* 10: 45-83.

Searle, J. 1969. *Speech Acts: An essay in the philosophy of language*. Cambridge, Engl.: Cambridge University Press.

Searle, J. 1975. Indirect Speech Acts. In P. Cole and J. Morgan, eds. *Syntax and Semantics Vol. 3, Speech Acts*, New York: Academic Press, pp. 59-82.

Sidner, C. L. 1979. *Towards a Computational Theory of Definite Anaphora Comprehension in English Discourse*. Ph. D. thesis. Technical report 537, Artificial Intelligence Laboratory, MIT, Cambridge, MA.

Sidner, C. L. 1981. Focusing for interpretation of pronouns. *Computational Linguistics* 7(4): 217-231.

Sidner, C. L. 1982. *Protocols of Users Manipulating Visually Presented Information with Natural Language*. Technical report, Bolt, Berenek and Newman, Cambridge, MA.

Sidner, C. L. 1983. What the speaker means: The recognition of speakers' plans in discourse. *International Journal of Computers and Mathematics, Special Issue in Computational Linguistics* 9(1): 71-82.

Sidner, C. L. 1985. Plan parsing for intended response recognition in discourse. *Computational Intelligence* 1(1): 1-10.

Silverman, K. 1987. *Natural Prosody for Synthetic Speech*. Ph. D. thesis, Cambridge University, Cambridge, Engl.

Stefik, M. 1981. Planning with constraints. *Artificial Intelligence* 16: 111-140.

Tate, A. 1984. Goal structure-capturing the intent of plans. In *Proceedings of the Sixth European Conference on Artificial Intelligence*, Pisa, Italy, pp. 273-276.

van Dijk, T. A. 1972. *Some Aspects of Text Grammars*. The Hague: Mouton.

Walker, D. 1978. *Understanding Spoken Language*. New York: Elsevier North-Holland.

Webber, B. L. 1980. *A Computational Approach to Discourse Anaphora*. New York: Garland.

Webber, B. L. 1983. So what can we talk about now. In M. Brady, ed. *Computational Approaches to Discourse*. Cambridge, MA: MIT Press.

Webber, B. L. 1987. The interpretation of tense in discourse. In *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics*, Stanford, CA, pp. 147-154.

Wilensky, R. 1983. *Planning and Understanding*. Reading, MA: Addison-Wesley.

Wilkins, D. E. 1984. Domain-independent planning: Representation and plan generation. *Artificial Intelligence* 22: 269-301.

Wilkins, D. E. 1985. *Hierarchical Planning: Definition and Implementation*. Technical report 370, Artificial Intelligence Center, SRI International, Menlo Park, CA.

Winograd, T. 1971. *Procedures as a Representation for Data in a Computer Program for Understanding Natural Language*. Ph. D. thesis, Report no. TR-84, Project MAC, MIT, Cambridge, MA.

Winograd, T. 1978. *Understanding Natural Language*. New York: Academic Press.

Woods, W. A. 1978. Semantics and quantification in natural language question answering. In M. Yovits, ed. *Advances in Computers*. New York: Academic Press, pp. 1-87.

Woods, W. A., Kaplan, R. M., and Nash-Webber, B. L. 1972. *The Lunar Sciences Natural Language Information System: Final Report*. BBN Report 2378, Bolt, Beranek and Newman, Cambridge, MA.